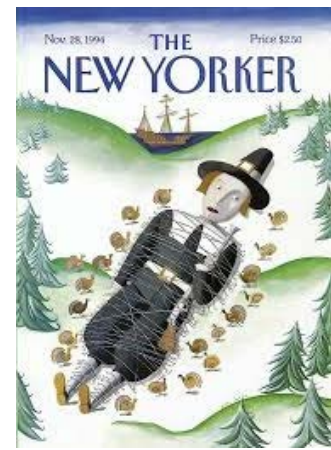


# Curveball

Stephen Jay Gould

*The New Yorker*, November 28, 1994



*The Bell Curve*, by Richard J. Herrnstein and Charles Murray (Free Press; \$30), subtitled *Intelligence and Class Structure in American Life*, provides a superb and unusual opportunity to gain insight into the meaning of experiment as a method in science. The primary desideratum in all experiments is reduction of confusing variables: we bring all the buzzing and blooming confusion of the external world into our laboratories and, holding all else constant in our artificial simplicity, try to vary just one potential factor at a time. But many subjects defy the use of such an experimental method—particularly most social phenomena—because importation into the laboratory destroys the subject of the investigation, and then we must yearn for simplifying guides in nature. If the external world occasionally obliges by holding some crucial factors constant for us, we can only offer thanks for this natural boost to understanding.

So, when a book garners as much attention as *The Bell Curve*, we wish to know the causes. One might suspect the content itself—a startlingly new idea, or an old suspicion newly verified by persuasive data—but the reason might also be social acceptability, or even just plain hype. *The Bell Curve*, with its claims and supposed documentation that race and class differences are largely caused by genetic factors and are therefore essentially immutable, contains no new arguments and presents no compelling data to support its anachronistic social Darwinism, so I can only conclude that its success in winning attention must reflect the depressing temper of our time—a historical moment of unprecedented ungenerosity, when a mood for slashing social programs can be powerfully abetted by an argument that beneficiaries cannot be helped, owing to inborn cognitive limits expressed as low IQ scores.

*The Bell Curve* rests on two distinctly different but sequential arguments, which together encompass the classic corpus of biological determinism as a social philosophy. The first argument rehashes the tenets of social Darwinism as it was originally constituted. "Social Darwinism" has often been used as a general term for any evolutionary argument about the biological basis of human differences, but the initial nineteenth-century meaning referred to a specific theory of class stratification with industrial societies, and particularly to the idea that there was a permanently poor underclass consisting of genetically inferior people who had precipitated down into their inevitable fate. The theory arose from a paradox of egalitarianism: as long as people remain on top of the social heap by accident of a noble name or parental wealth, and as long as members of despised castes cannot rise no matter what their talents, social stratification will not reflect intellectual merit, and brilliance will be distributed across all classes; but when true equality of opportunity is attained smart people rise and the lower classes become rigid, retaining only the intellectually incompetent.

This argument has attracted a variety of twentieth-century champions, including the Stanford psychologist Lewis M. Terman, who imported Alfred Binet's original test from France, developed the Stanford-Binet IQ test, and gave a hereditarian interpretation to the results (one that Binet had vigorously rejected in developing this style of test); Prime Minister Lee Kuan Yew of Singapore, who tried to institute a eugenics program of rewarding well-educated women for higher birth rates; and Richard Herrnstein, a co-author of *The Bell Curve* and also the author of a 1971 *Atlantic Monthly* article that presented the same argument without the documentation. The general claim is neither uninteresting nor illogical, but it does require the validity of four shaky premises, all asserted (but hardly discussed or defended) by Herrnstein and Murray. Intelligence, in their formulation, must be depictable as a single number, capable of ranking people in linear order, genetically based, and effectively immutable. If any of these premises are false, their entire argument collapses. For example, if all are true except immutability, then programs for early intervention in education might work to boost IQ permanently, just as a pair of eyeglasses may correct a genetic defect in vision. The central argument of *The Bell Curve* fails because most of the premises are false.

Herrnstein and Murray's second claim, the lightning rod for most commentary extends the argument for innate cognitive stratification to a claim that racial differences in IQ are mostly determined by genetic causes—small difference for Asian superiority over Caucasian, but large for Caucasians over people of African descent. This argument is as old as the study of race, and is most surely fallacious. The last generation's discussion centered on Arthur Jensen's 1980 book *Bias in Mental Testing* (far more elaborate and varied than anything presented in *The Bell Curve*, and therefore still a better source for grasping the argument and its problems), and on the cranky advocacy of William Shockley, a Nobel Prize-winning physicist. The central fallacy in using the substantial heritability of within-group IQ (among whites, for example) as an explanation of average differences between groups (whites versus blacks, for example) is now well known and acknowledged by all, including Herrnstein and Murray, but deserves a restatement by example. Take a trait that is far more heritable than anyone has ever claimed IQ to be but is politically uncontroversial—body height. Suppose that I measured the heights of adult males in a poor Indian village beset with nutritional deprivation, and suppose the average height of adult males is five feet six inches. Heritability within the village is high, which is to say that tall fathers (they may average five feet eight inches) tend to have tall sons, while short fathers (five feet four inches on average) tend to have short sons. But this high heritability within the village does not mean that better nutrition might not raise average height to five feet ten inches in a few generations. Similarly, the well-documented fifteen-point average difference in IQ between blacks and whites in America, with substantial heritability of IQ in family lines within each group, permits no automatic conclusion that truly equal opportunity might not raise the black average enough to equal or surpass the white mean.

Disturbing as I find the anachronism of *The Bell Curve*, I am even more distressed by its pervasive disingenuousness. The authors omit facts, misuse statistical methods, and seem unwilling to admit the consequence of their own words.

The ocean of publicity that has engulfed *The Bell Curve* has a basis in what Murray and Herrnstein, in an article in *The New Republic* last month [Oct. 31, 1994], call "the flashpoint

of intelligence as a public topic: the question of genetic differences between the races." And yet, since the day of the book's publication, Murray (Herrnstein died a month before the book appeared) has been temporizing, and denying that race is an important subject in the book at all; he blames the press for unfairly fanning these particular flames. In *The New Republic* he and Herrnstein wrote, "Here is what we hope will be our contribution to the discussion. We put it in italics; if we could, we would put it in neon lights: *The answer doesn't much matter.*"

Fair enough, in the narrow sense that any individual may be a rarely brilliant member of an averagely dumb group (and therefore not subject to judgment by the group mean), but Murray cannot deny that *The Bell Curve* treats race as one of two major topics, with each given about equal space; nor can he pretend that strongly stated claims about group differences have no political impact in a society obsessed with the meanings and consequences of ethnicity. The very first sentence of *The Bell Curve's* preface acknowledges that the book treats the two subjects equally: "This book is about differences in intellectual capacity among people and groups and what those differences mean for America's future." And Murray and Herrnstein's *New Republic* article begins by identifying racial differences as the key subject of interest: "The private dialogue about race in America is far different from the public one."

Furthermore, Herrnstein and Murray know and acknowledge the critique of extending the substantial heritability of within-group IQ to explain differences between groups, so they must construct an admittedly circumstantial case for attributing most of the black-white mean difference to irrevocable genetics—while properly stressing that the average difference doesn't help in judging any particular person, because so many individual blacks score above the white mean in IQ. Quite apart from the rhetoric dubiety of this old ploy in a shopworn genre—"Some of my best friends are Group X"—Herrnstein and Murray violate fairness by converting a complex case that can yield only agnosticism into a biased brief for permanent and heritable difference. They impose this spin by turning every straw on their side into an oak, while mentioning but downplaying the strong circumstantial case for substantial malleability and little average genetic difference. This case includes such evidence as impressive IQ scores for poor black children adopted into affluent and intellectual homes; average IQ increases in some nations since the Second World War equal to the entire fifteen-point difference now separating blacks and whites in America; and failure to find any cognitive differences between two cohorts of children born out of wedlock to German women, reared in Germany as Germans, but fathered by black and white American soldiers.

*The Bell Curve* is even more disingenuous in its argument than in its obfuscation about race. The book is a rhetorical masterpiece of scientism, and it benefits from the particular kind of fear that numbers impose on nonprofessional commentators. It runs to 845 pages, including more than a hundred pages of appendixes filled with figures. So their text looks complicated, and reviewers shy away with a knee-jerk claim that, while they suspect fallacies of argument, they really cannot judge. In the same issue of *The New Republic* as Murray and Herrnstein's article, Mickey Kaus writes, "As a lay reader of 'The Bell Curve,' I am unable to judge fairly," and Leon Wieseltier adds, "Murray, too, is hiding the hardness of

his politics behind the hardness of his science. And his science, for all I know, is soft.... Or so I imagine. I am not a scientist. I know nothing about psychometrics." And Peter Passell, in the *Times*: "But this reviewer is not a biologist, and will leave the argument to experts."

The book is in fact extraordinarily one-dimensional. It makes no attempt to survey the range of available data, and pays astonishingly little attention to the rich and informative history of its contentious subject. (One can only recall Santayana's dictum now a cliché of intellectual life: "Those who cannot remember the past are condemned to repeat it.") Virtually all the analysis rests on a single technique applied to a single set of data—probably done in one computer run. (I do agree that the authors have used more appropriate technique and the best source of information. Still, claims as broad as those advanced in *The Bell Curve* simply cannot be properly defended—that is, either supported or denied—by such a restricted approach.) The blatant errors and inadequacies of *The Bell Curve* could be picked up by lay reviewers if only they would not let themselves be frightened by numbers—for Herrnstein and Murray do write clearly, and their mistakes are both patent and accessible.

While disclaiming his own ability to judge, Mickey Kaus, in *The New Republic*, does correctly identify the authors' first two claims that are absolutely essential "to make the pessimistic 'ethnic difference' argument work": "1) that there is a single, general measure of mental ability; 2) that the IQ tests that purport to measure this ability...aren't culturally biased."

Nothing in *The Bell Curve* angered me more than the authors' failure to supply any justification for their central claim, the sine qua non of their entire argument: that the number known as *g*, the celebrated "general factor" of intelligence, first identified by British psychologist Charles Spearman, in 1904, captures a real property in the head. Murray and Herrnstein simply declare that the issue has been decided, as in this passage from their *New Republic* article: "Among the experts, it is by now beyond much technical dispute that there is such a thing as a general factor of cognitive ability on which human beings differ and that this general factor is measured reasonably well by a variety of standardized tests, best of all by IQ tests designed for that purpose." Such a statement represents extraordinary obfuscation, achievable only if one takes "expert" to mean "that group of psychometricians working in the tradition of *g* and its avatar IQ" The authors even admit that there are three major schools of psychometric interpretation and that only one supports their view of *g* and IQ.

But this issue cannot be decided, or even understood, without discussing the key and only rationale that has maintained *g* since Spearman invented it: factor analysis. The fact that Herrnstein and Murray barely mention that factor-analytic argument forms a central indictment of *The Bell Curve* and is an illustration of its vacuousness. How can the authors base an 800-page book on a claim for the reality of IQ as measuring a genuine, and largely genetic, general cognitive ability—and then hardly discuss, either pro or con, the theoretical basis for their certainty?

Admittedly, factor analysis is a difficult mathematical subject, but it can be explained to lay readers with a geometrical formulation developed by L. L. Thurstone, an American psychologist, in the 1930s and used by me in a full chapter on factor analysis in my 1981

book *The Mismeasure of Man*. A few paragraphs cannot suffice for adequate explanation, so, although I offer some sketchy hints below, readers should not question their own IQs if the topic still seems arcane.

In brief, a person's performance on various mental tests tends to be positively correlated—that is, if you do well on one kind of test, you tend to do well on the other kinds. This is scarcely surprising, and is subject to interpretation that is either purely genetic (that an innate thing in the head boosts all performances); the positive correlations in themselves say nothing about causes. The results of these tests can be plotted on a multidimensional graph with an axis for each test. Spearman used factor analysis to find a single dimension—which he called *g*—that best identifies the common factor behind positive correlations among the tests. But Thurstone later showed that *g* could be made to disappear by simply rotating the dimensions to different positions. In one rotation Thurstone placed the dimensions near the most widely separated attributes among the tests, thus giving rise to the theory of multiple intelligences (verbal, mathematical, spatial, etc., with no overarching *g*). This theory (which I support) has been advocated by many prominent psychometricians, including J. P. Guilford, in the 1950s, and Howard Gardner today. In this perspective *g* cannot have inherent reality, for it emerges in one form of mathematical representation for correlations among tests and disappears (or greatly attenuates) in other forms, which are entirely equivalent in amount of information explained. In any case, you can't grasp the issue at all without a clear exposition of factor analysis—and *The Bell Curve* cops out on this central concept.

As for Kaus's second issue, cultural bias, the presentation of it in *The Bell Curve* matches Arthur Jensen's and that of other hereditarians, in confusing a technical (and proper) meaning of "bias" (I call it "S-bias," for "statistical") with the entirely different vernacular concept (I call it "V-bias") that provokes popular debate. All these authors swear up and down (and I agree with them completely) that the tests are not biased—in the statistician's definition. Lack of S-bias means that the same score, when it is achieved by members of different groups, predicts the same thing; that is, a black person and a white person with identical scores will have the same probabilities for doing anything that IQ is supposed to predict.

But V-bias, the source of public concern, embodies an entirely different issue, which, unfortunately, uses the same word. The public wants to know whether blacks average 85 and whites 100 because society treats blacks unfairly—that is, whether lower black scores record biases in this social sense. And this crucial question (to which we do not know the answer) cannot be addressed by a demonstration that S-bias doesn't exist, which is the only issues analyzed, however correctly, in *The Bell Curve*

The book is also suspect in its use of statistics. As I mentioned, virtually all its data derive from one analysis—a plotting, by a technique called multiple regression, of social behaviors that agitate us, such as crime, unemployment, and births out of wedlock (known as dependent variables), against both IQ and parental sociometric status (known as independent variables). The authors first hold IQ constant and consider the relationship of social behaviors to parental socioeconomic status. They then hold socioeconomic status constant and consider the relationship of the same social behaviors to IQ. In general, they find a higher correlation with IQ than with socioeconomic status; for example, people with

low IQ are more likely to drop out of high school than people whose parents have low socioeconomic status.

But such analyses must engage two issues—the form and the strength of the relationship—and Herrnstein and Murray discuss only the issue that seems to support their viewpoint, while virtually ignoring (and in one key passage almost willfully hiding) the other. Their numerous graphs present only the form of the relationships; that is, they draw the regression curves of their variables against IQ and parental socioeconomic status. But, in violation of all statistical norms that I've even learned, they plot only the regression curve and do not show the scatter of variation around the curve, so their graphs do not show anything about the strength of the relationships—that is, the amount of variation in social factors explained by IQ and socioeconomic status. Indeed, almost all their relationships are weak: very little of the variation in social factors is explained by either independent variable (though the form of this small amount of explanation does lie in their favored direction). In short, their own data indicate that IQ is not a major factor in determining variation in nearly all the social behaviors they study—and so their conclusions collapse, or at least become so greatly attenuated that their pessimism and conservative social agenda gain no significant support.

Herrnstein and Murray actually admit as much in one crucial passage, but then they hid the pattern. They write, "It [cognitive ability] almost always explains less than 20 percent of the variance, to use the statistician's term, usually less than 10 percent and often less than 5 percent. What this means in English is that you cannot predict what a given person will do from his IQ score.... On the other hand, despite the low association at the individual level, large differences in social behavior separate groups of people when the groups differ intellectually on the average." Despite this disclaimer, their remarkable next sentence makes a strong causal claim. "We will argue that intelligence itself, not just its correlation with socio-economic status, is responsible for these group differences." But a few percent of statistical determination is not causal explanation. And the case is even worse for their key genetic argument, since they claim a heritability of about 60 percent for IQ, so to isolate the strength of genetic determination by Herrnstein and Murray's own criteria you must nearly halve even the few percent they claim to explain.

My charge of disingenuousness receives its strongest affirmation in a sentence tucked away on the first page of Appendix 4, page 593: the authors state, "In the text, we do not refer to the usual measure of goodness of fit for multiple regressions,  $R^2$ , but they are presented here for the cross-sectional analyses." Now, why would they exclude from the text, and relegate to an appendix that very few people will read, or even consult, a number that, by their own admission, is "the usual measure of goodness of fit"? I can only conclude that they did not choose to admit in the main text the extreme weakness of their vaunted relationships.

Herrnstein and Murray's correlation coefficients are generally low enough by themselves to inspire lack of confidence. (Correlation coefficients measure the strength of linear relationships between variables; the positive values from 0.0 for no relationship to 1.0 for perfect linear relationship.) Although low figures are not atypical for large social-science surveys involving many variables, most of Herrnstein and Murray's correlations are very weak—often in the 0.2 to 0.4 range. Now, 0.4 may sound respectably strong, but—and this

is the key point— $R^2$  is the square of the correlation coefficient, and the square of a number between zero and one is less than the number itself, so a 0.4 correlation yields an  $R^2$  of only .16. In Appendix 4, then, one discovers that the vast majority of the conventional measures of  $R^2$ , excluded from the main body of the text, are less than 0.1.

These very low values of  $R^2$  expose the true weakness, in any meaningful vernacular sense, of nearly all the relationships that form the meat of *The Bell Curve*.

Like so many conservative ideologues who rail against the largely bogus ogre of suffocating political correctness, Herrnstein and Murray claim that they only want a hearing for unpopular views so that truth will out. And here, for once, I agree entirely. As a card-carrying First Amendment (near) absolutist, I applaud the publication of unpopular views that some people consider dangerous. I am delighted that *The Bell Curve* was written—so that its errors could be exposed, for Herrnstein and Murray are right to point out the difference between public and private agendas on race, and we must struggle to make an impact on the private agendas as well. But *The Bell Curve* is scarcely an academic treatise in social theory and population genetics. It is a manifesto of conservative ideology; the book's inadequate and biased treatment of data display its primary purpose—advocacy. The text evokes the dreary and scary drumbeat of claims associated with conservative think tanks: reduction or elimination of welfare, ending or sharply curtailing affirmative action in schools and workplaces, cutting back Head Start and other forms of preschool education, trimming programs for the slowest learners and applying those funds to the gifted. (I would love to see more attention paid to talented students, but not at this cruel price.)

The penultimate chapter presents an apocalyptic vision of a society with a growing underclass permanently mired in the inevitable sloth of their low IQs. They will take over our city centers, keep having illegitimate babies (for many are too stupid to practice birth control), and ultimately require a kind of custodial state, more to keep them in check—and out of high IQ neighborhoods—than to realize any hope of amelioration, which low IQ makes impossible in any case. Herrnstein and Murray actually write, "In short, by custodial state, we have in mind a high-tech and more lavish version of the Indian reservation for some substantial minority of the nation's population, while the rest of America tries to go about its business."

The final chapter tries to suggest an alternative, but I have never read anything more grotesquely inadequate. Herrnstein and Murray yearn romantically for the good old days of towns and neighborhoods where all people could be given tasks of value, and self-esteem could be found for people on all steps of the IQ hierarchy (so Forrest Gump might collect clothing for the church raffle, while Mr. Murray and the other bright ones do the planning and keep the accounts—they have forgotten about the town Jew and the dwellers on the other side of the tracks in many of these idyllic villages). I do believe in this concept of neighborhood, and I will fight for its return. I grew up in such a place in Queens. But can anyone seriously find solutions for (rather than important palliatives of) our social ills therein?

However, if Herrnstein and Murray are wrong, and IQ represents not an immutable thing in

the head, grading human beings on a single scale of general capacity with large numbers of custodial incompetents at the bottom, then the model that generates their gloomy vision collapses, and the wonderful variousness of human abilities, properly nurtured, reemerges. We must fight the doctrine of *The Bell Curve* both because it is wrong and because it will, if activated, cut off all possibility of proper nurturance for everyone's intelligence. Of course, we cannot all be rocket scientists or brain surgeons, but those who can't might be rock musicians or professional athletes (and gain far more social prestige and salary thereby), while others will indeed serve by standing and waiting.

I closed my chapter in *The Mismeasure of Man* on the unreality of *g* and the fallacy of regarding intelligence as a single-scaled, innate thing in the head with a marvelous quotation from John Stuart Mill, well worth repeating:

The tendency has always been strong to believe that whatever received a name must be an entity or being, having an independent existence of its own, and if no real entity answering to the name could be found, men did not for that reason suppose that none existed, but imagined that it was something particularly abstruse and mysterious.

How strange that we would let a single and false number divide us, when evolution has united all people in the recency of our common ancestry—thus undergirding with a shared humanity that infinite variety which custom can never stale. *E pluribus unum*.